

# Automatic determination of optimal spectral peaks for classification of Chinese tea varieties using laser-induced breakdown spectroscopy

Hongyang Zhang, Qibing Zhu\*, Min Huang, Ya Guo

(Key Laboratory of Advanced Process Control for Light Industry (Ministry of Education),  
Jiangnan University, Wuxi 214122, Jiangsu, China)

**Abstract:** The accurate identification of tea varieties is of great significance to ensure the interests of tea producers and consumers. As a non-destructive or micro damage detection method, laser-induced breakdown spectroscopy (LIBS) has been widely used in the quality detection or classification of agricultural products and food. The objective of this research was to automatically select optimal spectral peaks from the full LIBS spectra, and develop effective classification model for identifying tea varieties. The LIBS spectra covering the region 200-500 nm were measured for 600 Chinese tea leaves including six varieties (i.e. Longjing green tea, Jinhao black tea, Tie Guanyin, Huang Jinya, White peony tea, and Anhua dark tea). A total of 50 optimal spectral peaks were automatically selected from full LIBS spectra (6102) by using the uninformative variable elimination (UVE) and partial least squares projection analysis, and the selected spectral peaks mainly represent the elemental difference in C, Fe, Mg, Mn, Al and Ca. Partial Least Squares Discriminant Analysis (PLS-DA) was used for developing classification model using selected optimal spectral peaks, and yielded the 99.77% classification accuracy for 300 test samples was reached. The results indicate that the proposed method can be used to identify leaf varieties in various tea products.

**Keywords:** LIBS, tea varieties classification, feature selection, PLS projection algorithm, UVE, PLSDA

**DOI:** 10.25165/j.ijabe.20181103.3482

**Citation:** Zhang H Y, Zhu Q B, Huang M, Guo Y. Automatic determination of optimal spectral peaks for classification of Chinese tea varieties using laser-induced breakdown spectroscopy. Int J Agric & Biol Eng, 2018; 11(3): 154–158.

## 1 Introduction

Tea, containing amino acids, tea polyphenols, caffeine, and other chemical components and trace elements that are beneficial to the human body<sup>[1]</sup>, is one of the most frequently consumed beverages in the world. At present, about 18-20 billion cups of tea are consumed daily<sup>[2]</sup>, and the consumption demand is shown a growing trend accompanied by the increasing awareness of a healthy diet. Affected by tea plants, cultivation environment, processing technology, there are many tea varieties around the world. Different varieties have different flavor and nutritional elements, which determine the consumer's willingness to buy and price of tea. In order to obtain illegal benefits, the phenomenon that some vendors sell fake, low-quality tea has occurred frequently. Inauthentic tea products infringe the rights of broad consumers and disrupt the fair trade and international circulation of authentic tea products. Therefore, a method that can identify tea varieties quickly and accurately is important for the tea industry.

Currently, some methods, such as inductively coupled plasma emission spectroscopy<sup>[3]</sup>, inductively coupled plasma mass spectrometry<sup>[4]</sup>, atomic fluorescence spectrometry<sup>[5]</sup>, gas chromatography<sup>[6]</sup>, near-infrared spectroscopy<sup>[7]</sup>, and high

performance liquid chromatography<sup>[8]</sup> are used to analyze chemical compositions of tea leaves. However, although these methods can analyze tea composition, some methods require an external environment or complex sample preprocessing and do not have ability to remote control. Therefore, a fast and accurate classification method for tea species identification is of great significance.

In recent years, laser-induced breakdown spectroscopy (LIBS) has been widely used in environmental, industrial, and agricultural product detection, because it is rapid and accurate; in addition, this method is capable of micro-damage detecting and multiple elements simultaneously<sup>[9-12]</sup>. The LIBS technique produces several thousands to tens of thousands of variable (spectral) information, which is the basis for the quantitative/qualitative analysis of samples. However, the models developed using whole spectral lines of LIBS may not give the robust results due to the noise and redundancy containing in the whole LIBS spectra. Frank et al.<sup>[13]</sup> built several PLSDA models using different variable sets based on LIBS to identify 13 kinds of explosive sample. They found that the PLSDA model using the whole spectra can't obtain the most robust model. Therefore, an important problem using LIBS technique is how to select the representative spectral lines from whole spectra of LIBS data, and establish the identifying model using selected spectral lines. At present, the most used method for selecting representative spectral lines is based on the visual observation by operators. That is to say, by observing the intensity of the LIBS spectra, the spectral lines of some elements are selected from whole spectra of LIBS. For example, Yueh et al.<sup>[14]</sup> used hierarchical clustering analysis method, PLSDA, and artificial neural networks successively to classify biological tissues from different organs based on manually selected 21 spectral lines. Wang et al.<sup>[15]</sup> selected the spectral line intensities of CN, Ca, Al,

**Received date:** 2017-05-10 **Accepted date:** 2017-10-02

**Biographies:** **Hongyang Zhang**, Master, research interests: nondestructive detecting technology of agricultural products, Email: zhanghongyangwell@163.com; **Min Huang**, PhD, Professor, research interests: agricultural information acquisition and utilization, Email: huangmzqb@163.com; **Ya Guo**, PhD, Professor, research interests: agricultural information acquisition and utilization, Email: guoya68@163.com.

**\*Corresponding author:** **Qibing Zhu**, PhD, Professor, research interests: agricultural information acquisition and utilization. Jiangnan University, Wuxi 214122, Jiangsu, China. Tel: +86-15961753162, Email: zhuqib@163.com.

Mg, Mn, Fe, and C<sub>2</sub> as analysis indexes to identify six kinds of tea. However, this method relies on the experience of the observer, the selected spectral lines are often difficult to effectively reflect the difference of different samples, resulting to the poor classification accuracy.

There are many elements in tea leaves, the overlap of the spectral peaks for some elements is serious, and the intensities of the spectral peaks for trace elements are weak. Hence, it is a key problem how to select the appropriate spectral lines from the LIBS spectra to develop the accurate classification model for tea variety. In this study, uninformative variable elimination (UVE) and partial least squares projection algorithm were used to obtain the optimal LIBS spectral features of tea varieties. The noise in the LIBS spectrum has an effect on the spectral signal. UVE is a method for variable selection and evaluating the reliability of each variable in the model, i.e., the stability of each variable for noise, and eliminating the uninformative variables. Then, PLS projection algorithm can select a small amount of informative and representative characteristic variables and further reduce the redundancy characteristic variables. On the basis of the PLS projection algorithm, the tea spectral feature wavelengths were determined by PLS projection coefficients. Therefore, the objective of this study is to: 1) Select the optimal spectral peaks from the whole spectra of tea, using UVE and PLS projection algorithm; 2) Develop classification model for tea varieties based on selected optimal spectral peaks using PLSDA algorithm.

## 2 Materials and methods

### 2.1 Sample preparation

Six varieties of Chinese tea leaves, namely, Longjing green tea (LJ), Jinhao tea (JH), Tie Guanyin (TG), Huangjinya (HJ), White peony tea (WP), and Anhua black tea (AH), were analyzed in this study. These tea leaves were produced in 2016 and provided by the Institute of Tea in Wuxi, China. These tea leaves are popular and widely consumed in China. The physical properties of each sample, such as dryness, lapping uniformity, and density, have a great impact on spectral signal<sup>[16]</sup>. The samples were pretreated to improve the accuracy and stability of the results. The tea leaves were dried for 5 h at approximately 50°C using a drying box. The dried tea leaves were then ground separately into powder. Each resulting tea powder was filtered for sample uniformity. Finally, each tea powder was compressed to tablets (Φ15 mm×3 mm) at 25 MP for 3 min using a mechanical compression machine. For each type of tea, 100 samples were made. The experiment obtained a total of 600 samples for six varieties. Figure 1 shows the tea tablets of six varieties after handling.



Figure 1 Tea samples

### 2.2 Instrumental setup

The experimental setup of the LIBS is shown in Figure 2. A 190 mJ passively Q-seitched Nd: YAG laser (Quantel, French)

operating at 1064 nm with a pulse duration of 6.0 ns was used as an excitation light source. The laser was focused on the surface of the target samples using a plano-convex quartz lens (100 mm focal length) to generate intense plasma. Optical radiation was resolved by a three-channel spectrometer (AVS-DESKTOP-USB2, Avantes, Netherlands). The spectrometer detection zone was 200-500 nm, and the resolution was 0.08 nm. The detector is an intensified charge-coupled device (Andor, ICCD, 1024 × 1024 pixels). The built-in camera imaging system was connected to a computer. The data acquisition and analysis were conducted by using an online PC. DG535 (Stanford SRS delay/pulse signal generator) was used to control the delay time of the signal acquisition and platform rotation. The samples were placed on a rotating platform to ensure that the detection is uniform, stable, and not repeated.

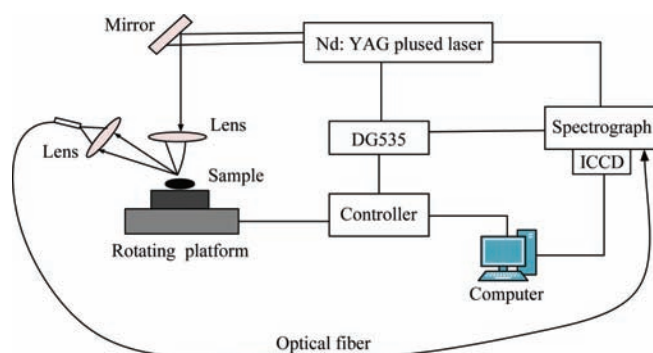


Figure 2 Schematic diagram of the experimental setup

LIBS signal is affected by laser energy, delay time, and gate width. The appropriate experimental parameters are thus important in obtaining stable spectral data. In this study, a series of preliminary experiments was carried out to reduce the interference from the noise and obtain a stable LIBS spectral signal. Furthermore, the following conditions were used: 100 mJ single laser pulse energy, 1.5 μs ICCD delay time, and 1.05 ms gate width. For each sample, the six repeated measurements were carried, and their average value was used as spectral data for further analysis.

### 2.3 Optimal spectral peaks selection

#### 2.3.1 Automatic determination of spectral peaks

The spectral peak of the LIBS spectrum represents the element information. In this experiment, the spectral peaks are automatically selected as follows:

Step 1: Given  $I_i(\lambda_j)$  is the mean intensity value of the  $i^{\text{th}}$  variety at wavelength  $\lambda_j$ . For three adjacent wavelengths  $\lambda_{j-1}, \lambda_j, \lambda_{j+1}$ , if  $I_i(\lambda_{j-1}) < I_i(\lambda_j) > I_i(\lambda_{j+1})$  is true, the wavelength  $\lambda_j$  is selected to represent the position of spectral peak for the  $i^{\text{th}}$  variety, and all of the selected wavelengths form a collection set  $Sel_i$ .

Step 2: For a given wavelength  $\lambda$ , if  $\lambda$  is contained by all of collection set  $Sel_i, i=1,2,\dots,L$ , the wavelength  $\lambda$  is finally selected to represent the position of spectral peak for all of varieties, and all of these selected wavelengths form a wavelength set  $F Sel$ .

Step 3: The spectral peaks for all of samples is obtained, which is intensity of samples in the wavelength set  $F Sel$ , and recorded as matrix  $X$ .

#### 2.3.2 Uninformative variable elimination

Due to influence of noise, some spectral peaks have high variance but small covariance with the dependent variable  $y$  (variety labels in this study), which may decrease the model performance. UVE is a method for variable selection and evaluating the reliability of each variable in the model. This method can choose the stable variables (spectral peaks) from original variables. The specific steps of the feature selection by

UVE method are as follows<sup>[17]</sup>:

In linear squares models, the Equation (1) can be used to compute the prediction:

$$\hat{y} = X\beta + b_0 \quad (1)$$

where,  $X$  is an  $z \times l$  matrix containing  $l$  spectral peaks of  $z$  samples;  $\beta$  is a  $l \times 1$  vector of regression coefficients and  $b_0$  is the model offset. A regression coefficient matrix  $\beta = [\beta_1, \dots, \beta_h]$  is calculated through a leave-one-out validation. Because each coefficient  $\beta_w$  represents the contribution of the corresponding variable to the established model, the reliability of each variable  $w$  can be quantitatively measured by the stability defined as:

$$s_w = \text{mean}(\beta_w) / \text{std}(\beta_w) \quad w = 1, \dots, l \quad (2)$$

where,  $\text{mean}(\beta_w)$  and  $\text{std}(\beta_w)$  are the mean and standard deviation of the regression coefficients of variable  $w$ . The variables whose stability is less than a threshold should be treated as uninformative and be eliminated.

In order to estimate a suitable cutoff threshold, an artificial random variable matrix  $N^{z \times l}$  with very small amplitude is added to the original data to compute stability. It is obvious that any variable whose stability is less than that of random variables should be known as uninformative and be eliminated. In practice, the cutoff threshold is generally defined by:

$$\text{cutoff} = k \times \max(\text{abs}(s_w)) \quad (3)$$

where,  $k$  is an regulative parameter, which determine the number of selected variables. In this work, the value is set to 0.8. Every original variable with equal or lower absolute value of  $\text{cutoff}$  is assumed to contain nothing but noise and is eliminated.

### 2.3.3 PLS projection analysis

In order to further reduce the redundancy of selected variables using UVE, the PLS projection analysis was applied in optimal variables selection after UVE operation. Because the PLS projection analysis takes into account the standard deviation of variables and their contribution to the dependent variable, the model developed using optimal variables selected by PLS algorithm has good prediction performance and robustness<sup>[18]</sup>. In this paper, the selected optimal variables with PLS projection are chosen to develop classification model.

### 2.4 PLSDA classification model

PLSDA belongs to group of the supervised learning methods, and is widely used for building linear discriminant models even with high dimensional data<sup>[19]</sup>. This method transforms the high dimensional data into a set of a few intermediate linear latent variables (components) that are useful for predicting the dependent variable. The dependent variable (class variable) in PLSDA model is a dummy variable that indicates whether a given sample belongs to a given class. The number of components has an important impact on model performance. In this paper, the number of components was determined by the leave-one-out cross-validation.

There are many elements in tea and its content is also different, which leads to the great intensity difference among different spectral peaks. A larger value of spectral peak would have a dominant predicting power over smaller value of spectral peak, even though the latter may be as important as, or even more important than, the former. To overcome the problem of large disparity in values among different spectral peaks, an auto scaling method was used before model developing, which is given as follows:

$$X_{i,j}^* = \frac{X_{i,j} - X_{\min}^j}{X_{\max}^j - X_{\min}^j} \quad (4)$$

where,  $X_{i,j}^*$  represents the rescaled  $j^{\text{th}}$  component for sample  $i$ ;  $X_{\min}^j$  and  $X_{\max}^j$  are the minimum and maximum value of the  $j^{\text{th}}$  component of all samples, respectively, and  $X_{i,j}$  denotes the original value of the  $j^{\text{th}}$  component for sample  $i$ . After min-max normalization, the data ranged from 0 to 1.

Prior to the model development, samples of each variety were randomly divided into two groups, 50% of the samples were used for training, and the remaining 50% samples were used for independent test. For a given analysis, the training and test results may vary depending on how the samples are actually selected. To better evaluate the performance of the PLSDA model for classifying the variety of tea, the above training and test procedures were run twenty times. Only averages over the 20 runs are reported in the following section.

All the algorithms were run in Matlab R2009b (The MathWorks Inc., USA) and PLS Toolbox 5.0 (Eigenvector Research, Inc., USA).

## 3 Results and discussion

### 3.1 Tea spectral characteristics

In this study, the spectrometer detection zone regions were from 200 nm to 500 nm. Figure 3 shows the LIBS spectrum of LJ. Its elemental compositions were identified and marked according to the standards of the National Institute of Standards and Technology (NIST)<sup>[20]</sup>. The characteristic spectral lines of C, Ca, Fe, Al, Mn, Mg, and Si were then observed. In addition, some molecular spectra, such as that of CN, occurred in the LIBS spectrum. The spectral lines of CN are caused by the CN diatomic molecules in the tea leaves. Others are caused by the reaction of the elemental C in the tea leaves with N in the air ( $C_2 + N_2 \rightarrow 2CN$ ). Time-resolved LIBS uses different decay times to distinguish CN molecular spectral lines<sup>[21]</sup>.

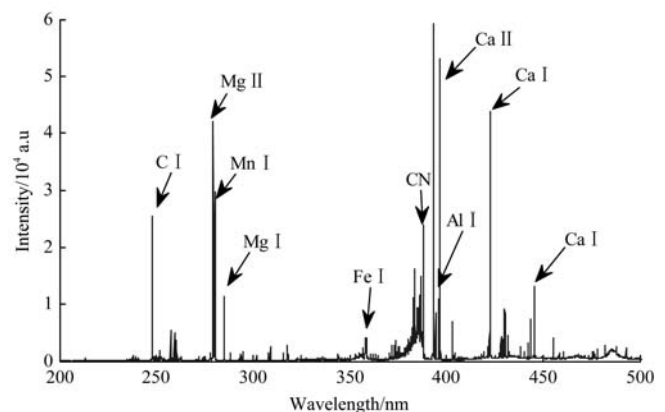


Figure 3 LIBS spectra of LJ green tea leaves at a wavelength range of 200-500 nm

The mean spectra of samples for each variety were shown in Figure 4. Although the spectra of the tea samples were similar, their LIBS spectral peak intensities varied. This finding indicates that the tea samples contain different elemental compositions.

### 3.2 Optimal spectral peaks selection

The full spectrum contains 6102 intensity values and contains a large number of redundant information and noise. According to the LIBS spectra, the 261 spectral peaks were selected as spectral features. In order to select some effective and anti-noise capability wavelengths, UVE algorithm is used to select wavelengths. In this study, 98 spectral lines were selected by UVE. Then, PLS projection analysis was used to select the features for the reduction of data redundancy and calculation

consumption of the model. According to the value of PLS projection coefficients, 50 characteristic spectral lines were

selected as features. Figure 5 shows the wavelength distribution selected by the PLS projection analysis algorithm.

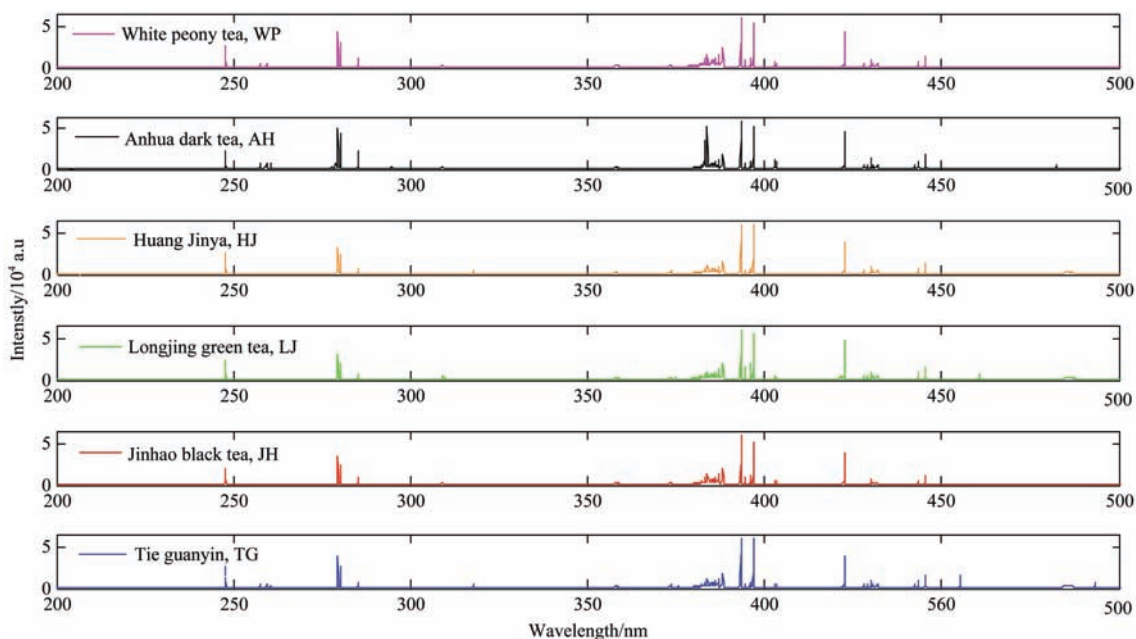


Figure 4 Mean spectra for each variety

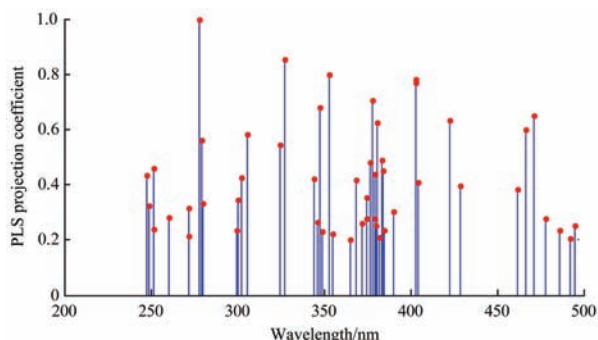


Figure 5 PLS projection coefficients of the optimal wavelengths

According to the NIST database, the spectral lines selected by PLS projection coefficient mainly represent C (247.85 nm), Fe (344.09 nm, 370.58 nm, 373.71 nm), Mg (279.54 nm, 280.26 nm, 383.81 nm), Mn (260.55 nm, 279.78 nm), Al (396.15 nm) and Ca (422.67 nm, 428.93 nm). Figure 6 shows the boxplots of intensity for some elements. Although there is a large variation of intensity for each variety, the some differences of intensity for different variety still can be observed. For example, AH and TG have a larger intensity value than the other varieties at 396.15 nm. These differences are a basis for classification and recognition of tea variety.

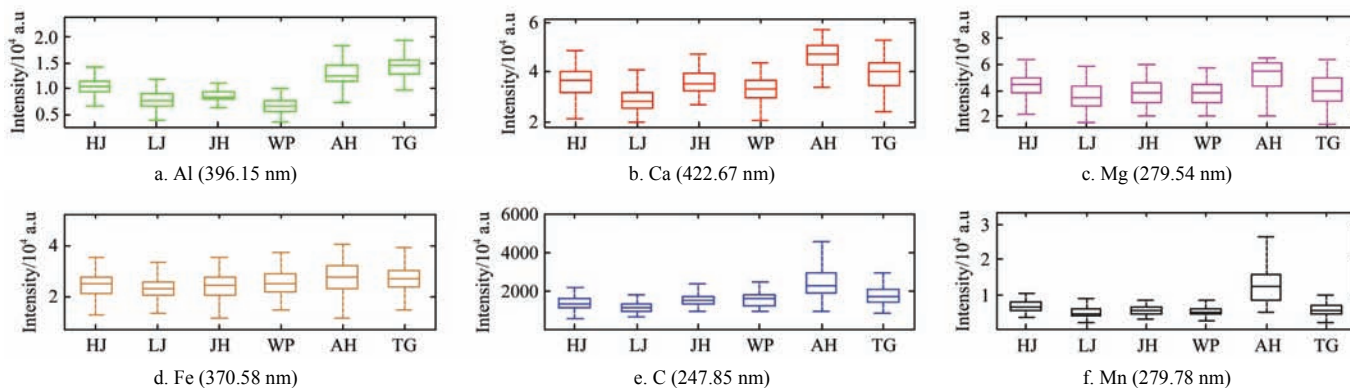


Figure 6 Intensity of different elements for six varieties of Chinese tea leaves

**3.3 Classification of the six varieties of Chinese tea leaves**

The classification results using PLSDA model based on different spectral features are present in Table 1. When the 261 spectral peaks were used for developing PLSDA model, the average correct identification accuracy was 97.67% for test set. The 98.67% classification accuracy was obtained by PLSDA model using the 98 spectral peaks selected by UVE. In comparison, the PLSDA model combined with UVE and the PLS projection analysis achieved the best classification results, with the 99.77% and 100% classification accuracies for the test set and training set, respectively. The high classification accuracy

obtained by UVE coupled with PLS projection analysis indicates that the selected spectral peaks can reflect the information difference containing in the original LIBS data. Table 2 shows detailed classification results for the test set of tea variety, using optimal spectral peaks selected by UVE and PLS projection analysis. The model achieved more than 99.5% classification accuracy for each variety.

The misclassification mainly occurred among the LJ and the AH samples, HJ and WP samples. This may be contributed to that there are similar spectral peaks among these varieties. Overall, the optimal spectral peaks selection eliminated invalid

information and sped up data storage and modeling. LIBS technique combined with PLSDA has advantages of speed and

accuracy. This study provides a useful method for element detection and discrimination of agricultural products.

**Table 1 Classification results of the six varieties of Chinese tea leaves using different PLSDA model**

| Wavelength selection (number of selected wavelength) | Training set/% | Test set/% |
|--|----------------|------------|
| Spectral peaks (261)                                 | 99.66          | 97.67      |
| UVE (98)   | 99.83          | 98.67      |
| UVE+PLS projection analysis (50)                     | 100            | 99.77      |

**Table 2 Confusion matrices for six varieties of Chinese tea leaves using PLSDA model combined with UVE and PLS projection analysis**

| Variety | Number <sup>[a]</sup> | Predicted variety <sup>[b]</sup> |            |            |            |            |            | False negative rate/% | False positive rate/% |
|---------|-----------------------|----------------------------------|------------|------------|------------|------------|------------|-----------------------|-----------------------|
|         |                       | HJ                               | LJ         | JH         | WP         | AH         | TG         |                       |                       |
| HJ      | 1000                  | 99.5 (995)                       | 0 (0)      | 0 (0)      | 0.4 (4)    | 0 (0)      | 0.1 (1)    | 0.5                   | 0.02                  |
| LJ      | 1000                  | 0 (0)                            | 99.5 (995) | 0 (0)      | 0 (0)      | 0.4 (4)    | 0.1 (1)    | 0.5                   | 0.06                  |
| JH      | 1000                  | 0 (0)                            | 0 (0)      | 100 (1000) | 0 (0)      | 0 (0)      | 0 (0)      | 0                     | 0                     |
| WP      | 1000                  | 0 (0)                            | 0.2 (2)    | 0 (0)      | 99.8 (998) | 0 (0)      | 0 (0)      | 0.2                   | 0.08                  |
| AH      | 1000                  | 0 (0)                            | 0.1 (1)    | 0 (0)      | 0 (0)      | 99.9 (999) | 0 (0)      | 0.1                   | 0.08                  |
| TG      | 1000                  | 0.1 (1)                          | 0 (0)      | 0 (0)      | 0 (0)      | 0 (0)      | 99.9 (999) | 0.1                   | 0.04                  |

Note: [a] The number is equal to the test sample number of each variety multiplied by 20 (20 runs).

[b] The number is equal to the predicted sample number of each variety multiplied by 20 (20 runs).

## 4 Conclusions

In this study, UVE and the PLS projection analysis were used to select the optimal spectral peaks automatically from LIBS data, for the identification of the Chinese tea varieties. The selected spectral lines represent the typical elements of C, Ca, Fe, Al, Mn, Mg, and Si in tea leaves. The PLSDA model based on selected spectral lines achieved the best performance, with 99.77% classification accuracy for test samples. This study provides a useful method for element detection and discrimination of agricultural products, different varieties of tea leaves in particular.

## Acknowledgments

This work was financially supported by the National Natural Science Foundation of China (Grant no. 61775086, 61772240), the Fundamental Research Funds for the Central Universities (JUSRP51730A).

## [References]

- Higdon J V, Frei B. Tea catechins and polyphenols: health effects, metabolism, and antioxidant functions. *Critical Reviews in Food Science and Nutrition*, 2003; 43(1): 89–143.
- Liao S, Kao Y H, Hiipakka R A. Green tea: biochemical and biological basis for health benefits. *Vitamins & Hormones*, 2001; 62(1): 1–94.
- Han L, Li R. Determination of minerals and trace elements in various tea by ICP-AES. *Spectroscopy & Spectral Analysis*, 2002; 22(2): 304–306.
- Wang X P, Ma Y J, Itoh M. Analysis of 23 mineral elements in tea samples collected from China and Japan by using ICP-AES and ICP-MS combined with a closed decomposition. *Spectroscopy and Spectral Analysis*, 2005; 25(10): 1703–1707.
- Cai L C. Arsenic speciation in drinking tea samples by hydride generation atomic fluorescence spectrometry. *Asian Journal of Chemistry*, 2013; 25(14): 8169–8172.
- Pongsuwan W, Bamba T, Yonetani T, Kobayashi A, Fukusaki E. Quality prediction of Japanese green tea using pyrolyzer coupled GC/MS based metabolic fingerprinting. *Journal of Agricultural & Food Chemistry*, 2008; 56(3): 744–750.
- Lee M S, Hwang Y S, Lee J, Choung M G. The characterization of caffeine and nine individual catechins in the leaves of green tea (*Camellia sinensis* L.) by near-infrared reflectance spectroscopy. *Food Chemistry*, 2014; 158(11): 351–357.
- He Q, Yao K, Jia D Y, Fan H J, Liao X P, Shi B. Determination of total catechins in tea extracts by HPLC and spectrophotometry. *Natural Product Research*, 2009; 23(1): 93–100.
- Zhang T, Wu S, Dong J, Li H. Quantitative and classification analysis of slag samples by laser-induced breakdown spectroscopy (LIBS) coupled with support vector machine (SVM) and partial least square (PLS) methods. *J. of Analytical Atomic Spectrometry*, 2015; 30(2): 368–374.
- Pandhija S, Rai A K. Screening of brick-kiln area soil for determination of heavy metal Pb using LIBS. *Environmental Monitoring & Assessment*, 2008; 148(1-4): 437–447.
- Multari R A, Cremers D A, Dupre J A, Gustafson J E. Detection of biological contaminants on foods and food surfaces using laser-induced breakdown spectroscopy (LIBS). *Journal of Agricultural & Food Chemistry*, 2013; 61(36): 8687–8694.
- Gottfried J L, Harmon R S, Jr F C D L, Miziolek A W. Multivariate analysis of laser-induced breakdown spectroscopy chemical signatures for geomaterial classification. *Spectrochimica Acta Part B Atomic Spectroscopy*, 2009; 64(10): 1009–1019.
- Jr L F C D, Gottfried J L. Influence of variable selection on partial least squares discriminant analysis models for explosive residue classification. *Spectrochimica Acta Part B Atomic Spectroscopy*, 2011; 66(2): 122–128.
- Yueh F Y, Zheng H, Singh J P, Burgess S. Preliminary evaluation of laser-induced breakdown spectroscopy for tissue classification. *Spectrochimica Acta Part B Atomic Spectroscopy*, 2009; 64(10): 1059–1067.
- Wang J M, Zheng P C, Liu H D, Fang L. Classification of Chinese tea leaves using laser-induced breakdown spectroscopy combined with the discriminant analysis method. *Analytical Methods*, 2016; 8(15): 3204–3209.
- Zheng P C, Shi M J, Wang J M, Liu H D. The spectral emission characteristics of laser induced plasma on tea samples. *Plasma Science and Technology*, 2015; 17(8): 664–670.
- Centner V, Massart D L, de Noord O E, De J S, Vandeginste B M, Sterna C. Elimination of uninformative variables for multivariate calibration. *Analytical Chemistry*, 1996; 68(21): 3851–3858.
- Dan, T N, Dai L K. Spectral wavelength selection based on PLS projection analysis. *Spectroscopy and Spectral Analysis*, 2009; 29(2): 351–354. (in Chinese)
- Barker M, Rayens W. Partial least squares for discrimination. *Journal of Chemometrics*, 2003; 17(3): 166–173.
- <http://www.physics.nist.gov/PhysRefData/ASD/lines-form.html>.
- Morel S, Leone N, Adam P, Amouroux J. Detection of bacteria by time-resolved laser-induced breakdown spectroscopy. *Applied Optics*, 2003; 42(30): 6184–6191.